

## Real Time Social Search and Related Problems

Ashish Goel, Stanford University

To answer search queries on a social network rich with user-generated content, it is desirable to give higher ranking to content that is closer to the individual issuing the query. To solve this problem exactly requires either prohibitively large pre-processing (e.g. maintaining a separate index of the entire corpus for every user) or prohibitively large effort for every query (e.g. a full breadth first search). In this talk, we will present an efficient approximate solution for this problem. Our solution requires maintaining a small number of indexes of the entire corpus. The pre-processing phase of our algorithm performs a small number (poly-logarithmic in  $N$ , where  $N$  is network size) of breadth-first search operations over the underlying social network, and is efficiently implementable offline. The indexing and querying operations can be performed efficiently (just two network calls in the typical scenario) in real-time on modern distributed stream processing platforms. The distance of the result returned is within an  $O(\log N)$ -factor of the closest result in the worst case, but our experimental evaluation shows that we typically find a closest result for realistic graphs. Time permitting, we will also describe an efficient distributed implementation of Locality Sensitive Hashing, which is another potentially useful tool in social search.

This is joint work with Bahman Bahmani, and appeared in WWW 2012.